Which roles & races does The Lord of the Rings series favor most?

Description of the data:

The data we use for this project comes from kaggle https://www.kaggle.com/mokosan/lord-of-the-rings-character-data. Two different datasets are used in our project: movies.csv and WordsbyCharacter.csv. The first one provides details about each movie in The Lord of the Rings series. The second one lists the total amount of words spoken by each character in each chapter* as well as the race of the character. We cleaned and aggregated the second dataset so that it finally has the amount of words spoken by each character in each movie*, and we also built a single csv file (Charecter.csv) to only store the race and name information for each character. As we only focus on The Lord of the Rings series, we ignored the data of the Hobbits. For the movie dataset, we kept each movie's runtime, budget, box revenue, awards and rating and used different scales in visualization to represent these information. For characters, we only care about their races and number of words they said in each movie. We can get some pieces of information out of the combined datasets and show it through our visualization, such as to find the character who plays the most important role in the movie series.

Mapping from data to visual elements:

Scales:

The movie part:

We use two circles to represent a single movie in the series.

The outer circle is used to show the box revenue of each movie, which is scaled by circle area. The larger the circle, the higher the box revenue.

The inner circle is used to show the budget of each movie, which is also scaled by circle area rather than radius. However, besides the fact that the domain and range used for outer and inner circle are both different, we tried to make the visualization more clear by making an internal tangent of two circles, which leads to a more obvious contrast in circle sizes than both centering them at the same point.
Also, we add one more feature to the outer circle which is the opacity, representing the Academy Awards winned by each movie. The opacity of inner circles are 0.8, however, for the outer one, the opacity scale is from 0.5 to 0.8 linearly, in a range that each circle can be displayed clearly and differently to others.

The character part:

For characters, as we care about the the amount of words spoken by each character and the character races, we put the characters as brands in the chord diagram. In the chord diagram, the circle is split into brands with the arc length of each group scaled to the total count of words spoken by each character.
By creating the connections between one movie and the number of words spoken by each character in it, the chords represent the relationships between characters and their weights in each movie. In sum, the total amount of words spoken by characters are scaled by the length of arcs on the circle using a linear scale, and the races of characters are scaled by colors. The chords from arcs to circles represent the amount of words by character in different movies.

Position:

With the logo centered in the top of the screen, we place the main visualization in the left, and the title as well as text for hove effects in the right. In order to let user see all the information in one page to form an idea what stories we are telling. We place the fieldset on the right side of the graph to show the related information when user have some interactions.

The legends are positioned in the bottom right corner with sizes that can be seen clearly but not too large to steal the attention.

The margin of the fieldset and legends are aligned to perform information in a tidy way. However, for the title part, as we want user to notice it, we decorated it in a bigger size and bolded it, and place it in a more center position.

Shape:

For legends, we use rectangles to represent the legends for races and circles to represent the legends for three outer and inner circles. Especially, as we have two properties (opacity and circle size) for outer circles, we made two individual legends with the same color to tell both the box revenue and winned awards information related to outer circles.

The texts for the hover effects are placed in the fieldset tag in the shape of a rectangle frame.

Color:

Different colors are picked for races so that characters in different races can be distinguished easily.

In addition, we tried to show colors in low saturation by choosing specific colors and setting the opacity of elements below 1. The purpose is to make the whole visualization look more comfortable and make users focus on the data itself rather than colors.

Hover effects:

We designed and implemented three types of hover effects for circles, individual chords and chord groups. A fieldset is used to display all the related information when user hovering on specific elements.

The group of chords and individual chords can both be highlighted under user selection and the corresponding information will be displayed on the right side in the fieldset.

The movie information, number of words spoken in each movie along with total words in the series are shown respectively when hovering on three circles and the arcs. For movie circle, the series name will be the legend value for the fieldset. For the character, the character's name is the legend value and the race of that character show in a race color scale. Also a small tooltip is put at the position of hovering for each arc to show the character name.

Story:

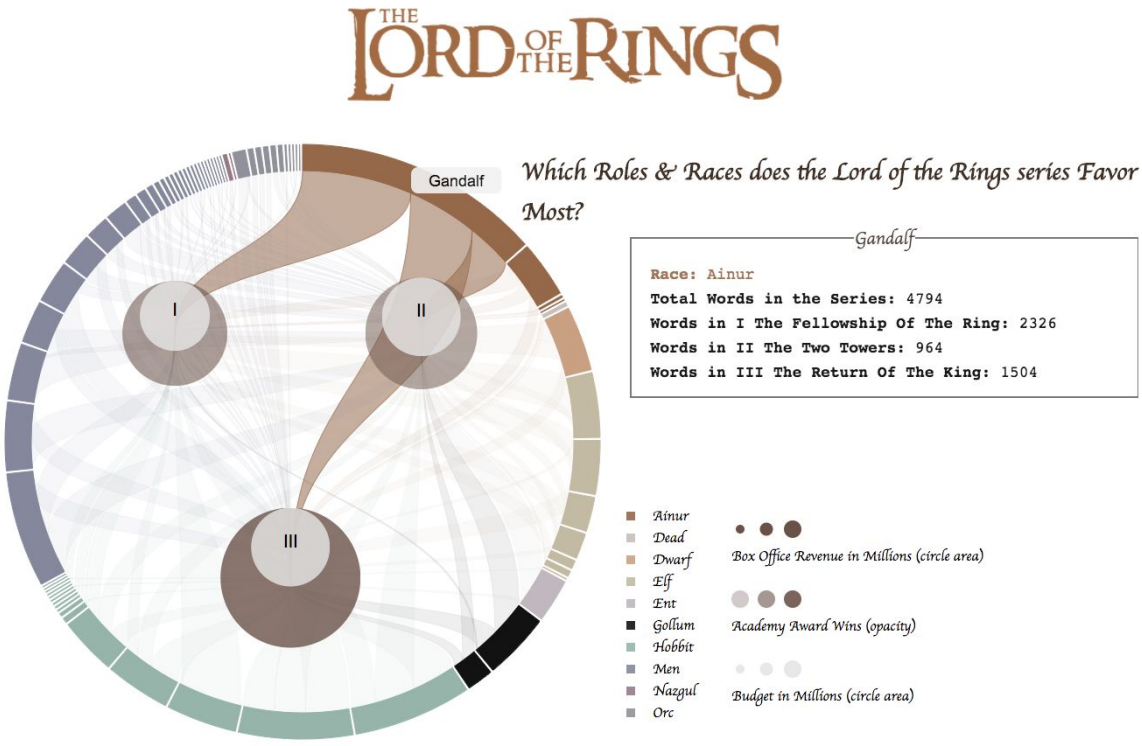
1. The preferences on races.
2. The main characters transferring from I to III.
3. Some races own only several important characters, some own lots of small characters.

For the movie part, we can clearly distinguish that the *Return of the King* is most successful in all aspects, like box revenue, awards winning.

From the race legend, we can tell that Race of men has the most total words. However, only several characters speak lots of words. Even the main character in the third series is Aragorn(the king of men race), Gandalf still the one who speak most of the words. As for Gandalf, he is the one who speak the largest number of words in the series, which make sense as he is the man who unfolding the whole story. Although there are not many Hobbits who show in the movies, Frodo and his friends and relatives still own a lot of words, which is true as Frodo is the one who had owned the ring. However, what surprised us is that Sam has more words than Frodo.

Although the results correspond to the our memory of the the movies we saw, there still some small points surprised us, like Gandalf says most the words, Legolas does not have much words than what we expected, Sam has more words than Frodo, and so on.

Result:



*The font may not be displayed normally on Windows